

Data Wrangling & Plotting

Student

7/7/2022

Taking a glimpse on the data:

```
## Rows: 3,548
## Columns: 7
## $ Success <int> 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, ~
## $ Height <int> 168, 162, 175, 169, 181, 173, 183, 173, 160, 190, 165, 188, ~
## $ Weight <int> 69, 99, 81, 75, 99, 74, 73, 83, 63, 96, 61, 82, 71, 83, 80, ~
## $ Age <int> 29, 54, 46, 68, 22, 20, 53, 65, 62, 52, 21, 65, 26, 39, 41, ~
## $ Age_Group <chr> "20-39", "40-59", "40-59", "60-79", "20-39", "20-39", "40-59~
## $ Gender <chr> "F", "M", "M", "F", "M", "M", "M", "M", "F", "M", "F", "M", ~
## $ Diabetes <int> 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, ~
```

```
##      Success Height Weight Age Age_Group Gender Diabetes
## 1         1    168     69  29    20-39      F          0
## 2         0    162     99  54    40-59      M          1
## 3         1    175     81  46    40-59      M          0
## 4         1    169     75  68    60-79      F          0
## 5         1    181     99  22    20-39      M          1
## 6         1    173     74  20    20-39      M          0
```

```
##
## No. of observations = 3548
##
##   Var. name obs. mean  median  s.d.   min.   max.
## 1 Success   3548 0.49    0       0.5    0      1
## 2 Height    3548 169.53 167     10.87  155    195
## 3 Weight     3548 80.87  79      13.3   60    110
## 4 Age        3548 50.07  51      17.6   20     80
## 5 Age_Group
## 6 Gender
## 7 Diabetes   3548 0.5     0       0.5    0      1
```

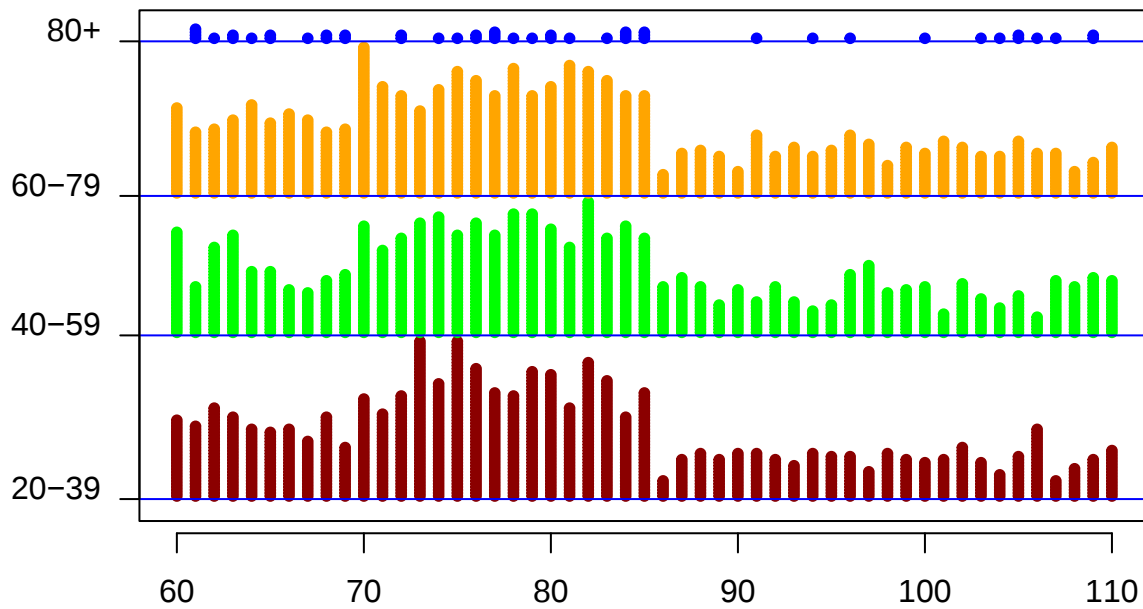
```
##
##      Cell Contents
## |-----|
## |              Count |
## | Chi-square contribution |
## |-----|
##
## Total Observations in Table:  3548
##
##              | data$Diabetes
## data$Age_Group |          0 |          1 | Row Total |
## -----|-----|-----|-----|
##          20-39 |        597 |        574 |        1171 |
##              |        0.088 |        0.089 |              |
## -----|-----|-----|-----|
##          40-59 |        573 |        566 |        1139 |
##              |        0.001 |        0.001 |              |
## -----|-----|-----|-----|
##          60-79 |        589 |        601 |        1190 |
##              |        0.179 |        0.182 |              |
## -----|-----|-----|-----|
##          80+ |         28 |         20 |         48 |
##              |        0.605 |        0.614 |              |
## -----|-----|-----|-----|
##      Column Total |        1787 |        1761 |        3548 |
## -----|-----|-----|-----|
##
##
```

```
## 'summarise()' has grouped output by 'Gender', 'Age_Group'. You can override
## using the '.groups' argument.
```

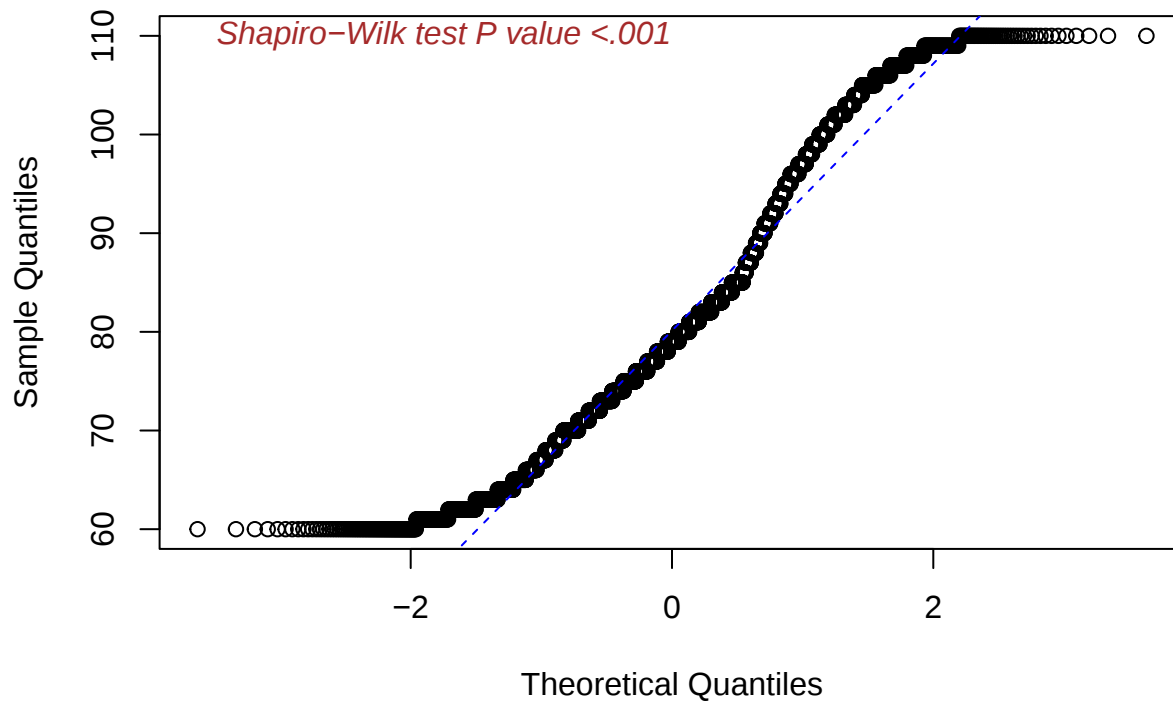
```
## # A tibble: 16 x 6
## # Groups:   Gender, Age_Group [8]
##   Gender Age_Group Diabetes      n avg_height median_weight
##   <chr>   <chr>      <int> <int>      <dbl>      <dbl>
## 1 F      20-39          0   322      162.         73
## 2 F      20-39          1   283      162.         73
## 3 F      40-59          0   272      162.         72
## 4 F      40-59          1   285      162.         73
## 5 F      60-79          0   301      162.         72
## 6 F      60-79          1   311      163.         72
## 7 F      80+           0    15      162.         69
## 8 F      80+           1    14      163         69.5
## 9 M      20-39          0   275      178.         89
## 10 M     20-39          1   291      178.         89
## 11 M     40-59          0   301      177.         90
## 12 M     40-59          1   281      176.         90
## 13 M     60-79          0   288      177.         92
## 14 M     60-79          1   290      177.         89.5
## 15 M     80+           0    13      175.        100
## 16 M     80+           1     6      181         93.5
```

Plotting and Graphics:

Distribution of data\$Weight by data\$Age_Group



Normal Q-Q plot of data\$Weight



```
## Warning in names(output)[names(output) == "mean.x"] <- paste("mean.",  
## as.character(substitute(x)), : number of items to replace is not a multiple of  
## replacement length
```

Mean and 95% CI of data\$Weight by data\$Height and data\$Diabetes

